

琉球大学学術リポジトリ

電子沖縄語辞典

メタデータ	言語: 出版者: 琉球大学工学部 公開日: 2007-08-23 キーワード (Ja): キーワード (En): Okinawan dialect, Dictionary, Computer, Window system, Conditional search 作成者: 高良, 富夫, 比嘉, 一秀, 比嘉, 智明, Takara, Tomio, Higa, Kazuhide, Higa, Tomoaki メールアドレス: 所属:
URL	http://hdl.handle.net/20.500.12000/1463

電子沖縄語辞典

高良富夫* 比嘉一秀** 比嘉智明***

An Electronic Okinawan Dialect Dictionary

Tomio TAKARA*, Kazuhide HIGA**, and Tomoaki HIGA***

Abstract

In the study of linguistics, we often use dictionaries not only to confirm the meaning of words, but also to investigate the phonemic law of a language. Also, for the study of the Ryukyuan dialects, we have used Ryukyuan dialect dictionaries many times. However, it is a very inefficient procedure to search for words and to process the data because we must do these all manually.

In order to facilitate the study of Ryukyuan dialects, we developed a software system which we named the Electronic Okinawan Dialect Dictionary; it is based on the dictionary of the Okinawa-Shuri dialect called "Okinawan Dialect Dictionary". Because the system is implemented on a personal computer, we can rapidly search for words and process the data thus obtained on the same computer. The remarkable features of this system are (1) easy operation using a window system, (2) rapid search, and (3) the use of the conditional search for phonemes. Part of the system had already been implemented in a speech synthesis system of the Shuri dialect for automatically detecting the accent type of a word.

Key Words: Okinawan dialect, Dictionary, Computer, Window system, Conditional search

1. まえがき

言語の研究においては、単語の意味を確認したり音韻の規則性を調べるため、辞典を多く活用する。琉球方言の研究においても方言辞典は活用されてきたが、これまで辞典の検索やそれを利用した統計処理などは全て手作業で行わなければならない、研究の能率は悪いものであった。

そこで本研究では、琉球方言の研究支援を目的として、「沖縄語辞典」[1]をもとに、コンピュータで辞書を高速に検索する「電子沖縄語辞典」を開発した。

開発にあたっては、方言の自動翻訳や文音声の合成などへ利用できるよう、今後の拡張性を考慮して、ソフトウェアは全て自作した。またウインドウ・システムを採用して操作性をよくし、TRIE構造を採用して検索を高速化した。

2. 電子沖縄語辞典の概要

本システムは、「沖縄語辞典」の本文編（沖縄首里方言辞典）および索引編（標準語引き）をコンピュータで検索できるようにしたものである。本システムで

受理：1993年11月8日

* 工学部情報工学科 Dept. of Information Engineering., Fac. of Eng.

** 日本アイビーエム(株) IBM Japan Ltd.

*** 沖縄日本電気ソフトウェア(株) Okinawa NEC Software Co.
沖縄言語研究センター研究発表会において1993年6月19日に発表済み

は、前者を検索することを琉和検索、後者を検索することを和琉検索と呼んでいる。本文編には、13,508語の首里方言の単語が登録されている。見出し語は、音韻記号（ローマ字）で表記されている。また見出し語の全てに、アクセント記号（平板型：0，下降型：1），品詞名およびその語の解説（意味や例文）が記されている。索引編の見出し語は平仮名で表記されており，その語に対応する首里方言が記されている。本システムでは，検索語をローマ字表記または平仮名表記で入力すると，自動的に，それぞれ琉和検索，和琉検索が実行される。「沖縄語辞典」で使用されている特殊な音韻記号は，本システムでは，簡単のため，表1のように置き換えている。

表1. 特殊文字の扱い

沖縄語辞典	本システム
ç	C
§	S
z	Z
?	?
]	"

本システムは，パーソナルコンピュータPC9801RAの上に構成されており，入出力装置としてマウスとプリンタを使用する。マウスは必須である。本システムを動作させるために必要な記憶容量は，約4.1Mバイトであるので，このコンピュータで動作させるためには，ハードディスクなどの外部記憶装置が必要である。本研究室では，外部記憶装置として光磁気ディスクを使用している。また，音声出力を使用する場合は，テキスト音声合成器「しゃべりん坊」[2]が必要である。

本システムのソフトウェア構成を図1に示す。本システムは，検索部，辞書操作部，出力部，条件検索部，OSシェルの各サブシステムからなる。メイン・メニューのサブシステム名をマウスで選択すると，そのサブシステムへ移行する。

検索部は，本システムの核をなす部分であり，琉和検索および和琉検索を行うことができる。検索語を指定する方法は3種類あり，検索語をキーボードから入

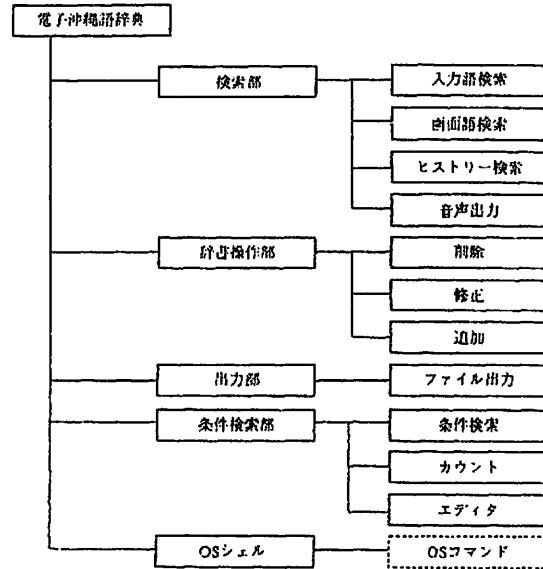


図1. 電子沖縄語辞典の構成

力する入力語検索，カレント・ウィンドウの文字列をマウス・カーソルで指定する画面語検索，検索した見出し語をヒストリーとしてシステムが記憶しておき，その中から見出し語をひとつ選択するヒストリー検索からなる。また音声合成器を利用し，カレント・ウィンドウの見出し語とその解説文を音声として出力する音声出力がある。

辞書操作部では，本システムに登録されている見出し語とその解説文の削除，修正，および登録されていない見出し語とその解説文の追加を行う。

出力部では，本システムに登録されている見出し語とその解説文をファイルに出力する。

条件検索部では，音韻表記上のある条件にあった見出し語をすべて取り出す条件検索，および任意の文字列を入力するとそれが現れる全ての見出し語を取り出し，その出現個数を数えるカウントを行う。また，検索された語を編集するエディタがある。

OSシェルでは，パーソナルコンピュータのMS-DOSコマンドを実行する。これにより，他のソフトウェアを使用して琉球方言の音声を出力することもできる。

以下に，各サブシステムの内容および使用法を示す。

3. 電子沖縄語辞典マニュアル

3.1 システムの起動

本システムを立ち上げるためには，実行形式のシス

テム本体と4個の辞書ファイルが同一ディレクトリ上にある必要がある。また、辞書操作部および条件検索部ではエディタとしてREDを使用するので、REDにパスを設定しておく。またデバイスとして、atok6.sys, atok6b.sys, mouse.sysを組み込んでおく。

この条件で

okinawa

と入力するとシステムが起動し、タイトル画面が現れる。ここでマウスを左クリックすると、メイン・メニューが現れる。

メイン・メニューの中の<検索>、<辞書操作>、<出力>、<条件検索>、<OS-Shell>、<終了>のいずれかにマウス・カーソルを合わせ左クリックし、メニューを選択する。これにより、指定したサブシステムに移行する。

3.2 検索部

ウィンドウの上下枠とタイトル枠以外の所でマウスを左クリックすると、<入力語>、<画面語>、<ヒストリー>、<音声出力>と表示されたウィンドウ(サブ・メニュー)が現れる。それらの中からひとつをマウスを左クリックすることにより選択する。なお、検索部の初期状態は入力語検索の平仮名入力となっている。

検索部では、検索した内容を画面に表示するため、マルチ・ウィンドウを採用している。まずマルチ・ウィンドウの基本操作法を示し、次に、それぞれの検索の使用法、および<音声出力>について述べる。

【マルチ・ウィンドウの基本操作】

希望する場所にマウスカーソルを移動させ左クリックして、ウィンドウの左上隅を決定する。すると、マウス・カーソルの移動に連れて、矩型の枠が大きくなったり、小さくなったりするので、枠の大きさを決定し、左クリックする。これにより、任意の大きさのウィンドウが作成される。

カレント・ウィンドウ(いま注目しているウィンドウ)には、3つのアイコン(印)がついており、それを左クリックすることにより以下の操作を行うことができる。

▲ 上スクロール：ウィンドウの表示内容を上に1行移動する。

▼ 下スクロール：同様に、下に1行移動する。

□ 閉じる：カレントウィンドウを消す。

【入力語検索】

画面下部に「検索語は? :」と表示されるので、検索したい語をキーボードにより入力する。琉和検索を行いたい場合は、ローマ字で入力し、和琉検索の場合は、平仮名で入力する。ローマ字入力と平仮名入力の切り替えは、《NFER》キーにより行う。

検索語が見つかり「表示ウィンドウを指定して下さい」と表示されるので、上記の【マルチ・ウィンドウの基本操作】に従ってウィンドウを開くと、検索された内容がウィンドウに表示される。検索語が本システムに登録されていない場合は、「……は見つかりません」と表示される。

【画面語検索】

画面下部に「検索語を指定して下さい」と表示されるので、カレント・ウィンドウの中の検索したい語を以下のようにして指定する。

(1) 語の先頭にマウス・カーソルを合わせ、左ボタンを押す。

(2) 左ボタンを押したまま語の末尾までマウス・カーソルを移動させボタンを離す。指定された語は、反転表示となる。

(3) 指定が終わったら左クリックする。指定を取り消したい場合は右クリックする。

検索語が検索されたら、【入力語検索】と同様にし、表示ウィンドウを開く。検索語が見つからない場合は、その旨のメッセージが表示される。

【ヒストリー検索】

画面左部にヒストリー・ウィンドウが現れ、これまでに検索された語が表示される。検索したい語にマウス・カーソルを合わせ左クリックする。【入力語検索】と同様にし、表示ウィンドウを開く。ここで検索を行わない場合は、右クリックする。

【音声出力】

これを選択すると、カレント・ウィンドウの内容が音声出力される。現在のところ、ローマ字(首里方言の音韻表記も)はアルファベットとして読みあげられる。

3.3 辞書操作部

辞書操作部では、<削除>、<修正>、<追加>と表示されたサブ・メニューが現れる。それらの中から

ひとつをマウスで選択すると、以下の操作を行うことができる。

【削除】

画面中央に「削除したい語は：」と表示されるので、語をキーボードにより入力する。その語が見つかった場合は「……を削除してよろしいですか(y/n)」と表示されるので、yまたはnをキーボードにより入力する。その後、削除されたかどうかのメッセージが表示される。

【修正】

画面中央に「修正したい語は：」と表示されるので、語をキーボードにより入力する。その語が見つかった場合はエディタREDが立ち上がり、修正したい語の解説文が表示される。REDを用いてその内容を修正する。修正を終了するためには、キーボードのf・1を押しくSave>を選択する。メイン・メニューが現れ、修正されたかどうかのメッセージが表示される。

【追加】

画面中央に「追加したいファイル名を入力して下さい：」と表示されるので、既存のファイル（辞書の内容が書き込まれているファイル）の追加を行う場合は、そのファイル名を入力し、新たにファイルを作成する場合は、新しいファイル名を入力する。新たにファイルを作成する場合および既存のファイルが見つからない場合は「ファイルを作成しますか？(y/n)」と表示されるので、yまたはnを入力する。

yを入力した場合は【修正】と同様、REDが立ち上がるので、それによりファイルを作成する。

メインメニュー画面が現れ、追加されたかどうかメッセージが表示される。

3.4 出力部

現在のところ、ファイルへの出力を行うことができる。画面中央に「ファイルに保存したい語は？：」と表示されるので、語をキーボードにより入力する。その語が見つかったら、「ファイル名（8文字以内）は？：」と表示されるので、ファイル名を入力する。保存が完了するとその旨メッセージが表示される。保存したい語が見つからない場合には、「……は見つかりません」と表示される。

3.5 条件検索部

<条件検索>、<カウント>、<エディタ>と表示されたサブ・メニューが現れるので、マウスにより選択し以下の操作を行う。なお、<条件検索>と<カウント>では、入力はローマ字で行う。従って、琉和検索の機能だけがある。

【条件検索】

ローマ字とワイルドカード*および\$を組み合わせて検索語を構成すると、その条件にあった見出し語をすべて取り出す。ここで、*は任意の文字列（0個以上）と一致することを表す。また、\$は任意の1個の文字と一致することを表す。例えば、以下のとおりである。

*： 全ての見出し語

mui： muiが含まれる全ての見出し語

\$\$\$： 3文字からなる全ての見出し語

\$\$\$\$mui： muiの前に4文字ある見出し語

*\$\$\$\$mui： muiの前に4文字以上ある見出し語

これにより任意の条件付き文字列を取り出すことができる。

使用法は以下のとおりである。まず、画面下部に「検索するファイル名は？：」と表示されるので、全ての首里方言から検索するのであれば、"outdt.txt"と入力する。もし、その前の<条件検索>で出力したファイルから検索するのであれば、そのファイル名を入力する。つぎに、「検索語は？：」と表示されるので、上記の規則に従い検索語を入力する。つぎに、「出力するファイル名は？：」と表示されるので、ファイル名を入力する。検索語の条件を満たす語が見つかったら「マッチした単語数=……」と表示される。条件を満たす語がない場合は、「マッチする単語はありません」と表示される。

【カウント】

任意の文字列を検索語として入力すると、その文字列を含む語をすべて取り出し、その文字列の出現回数を答える。ひとつの単語にその文字列が複数回現れる場合は、複数回として扱う。

使用法は上記の【条件検索】と同様である。ただし、ここでは上記のワイルドカードは使わない。また終了後に出現回数が表示されることが異なる。【条件検索】および【カウント】で出力されたファイルは、相互に入力ファイルとして使用できる。

【エディタ】

【条件検索】および【カウント】で出力したファイルをREDを用いて編集したり検索したりする。画面下部に「ファイル名は? :」と表示されるので、編集するファイルの名前を入力する。ここでREDが立ち上がり、そのファイルを操作することができる。

3.6 OSシェル部

OSへ抜けて、MS-DOSのコマンドおよび他のプログラムを実行する。使用方法は、まずメイン・メニューの<OSシェル>をマウスで選択する。これで、MS-DOSのコマンド・ラインでコマンドおよびプログラムが実行できる状態になる。本システムに戻るためには、“exit”と入力する。

4. 沖縄語辞典における母音の出現数

本システムの条件検索部を用いて、沖縄語辞典の出し語すべてに対して母音の出現数を調べた。結果を表2および表3に示す。

表2. 短母音の出現数

/a/	/i/	/u/	/e/	/o/	計
12,616	13,079	12,360	91	46	38,192
33.0%	34.1%	32.4%	0.2%	0.1%	100%

表3. 長母音の出現数

/aa/	/ii/	/uu/	/ee/	/oo/	計
2,425	2,118	1,590	1,721	1,303	9,157
26.5%	23.1%	17.4%	18.8%	14.2%	100%

この結果を伊波の調査[3]と比較すると、/u/と/ee/がわずかに少な目となっているが、出現頻度とその順位ともにはば一致している。ただし、本調査は単語単位で調査したものであるが、伊波のそれは文章中の出現頻度である。また、本調査の音素の総数は、伊波のその約10倍である。

5. むすび

琉球方言の研究を支援するシステムとして、「沖縄語辞典」をコンピュータで検索する「電子沖縄語辞典」を構成した。本システムの特長は、ウインドウ・システムにより操作性がよいこと、検索が速いこと、音韻の条件付き検索が行えることである。本システムの一部は、すでに首里方言のテキスト音声合成システム[4]においてアクセント型を自動的に決定するために使用されている。

今後の課題としては、以下のことが挙げられる。

- (1) 首里方言のテキスト音声合成システムを応用し、ウインドウの内容を方言を交えて読みあげることができるようにする。
- (2) 条件検索の条件として、AND, OR, 否定などを追加し、さらに有用なものにする。
- (3) これを琉球方言の他の方言にも応用できるものにする。
- (4) コンピュータ・ネットワークを介して他のコンピュータからも利用できるようにする。
- (5) 琉球方言の自動翻訳に応用する。

謝辞：本研究の一部は東京倶楽部の研究助成金によります。記して感謝致します。

文 献

- [1] 国立国語研究所編：“沖縄語辞典”，大蔵省印刷局（昭和58-04）。
- [2] NTTデータ通信（株）：“「しゃべりん坊」操作マニュアル”，（平成1-03）。
- [3] 大藤時夫，外間守善：“金城朝永全集 上巻 言語・文学編”，沖縄タイムス，pp.33-35（昭和49-01）。
- [4] 竹原康一：“琉球方言のテキスト音声合成システム”，琉球大学大学院工学研究科修士論文（平成3-03）。